Virology 447 (2013) 300–311

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/yviro

Diversity of human papillomaviruses in skin lesions

Johanna Ekström [a,b], Laila Sara Arroyo Mühr [b], Davit Bzhalava [b], Anna Söderlund-Strand [c], Emilie Hultin [b], Peter Nordin [d], Bo Stenquist [e], John Paoli [e], Ola Forslund [a], Joakim Dillner [b,*]

[a] Department of Laboratory Medicine, Lund University, SE 20502 Malmö, Sweden
[b] Departments of Laboratory Medicine, Medical Epidemiology & Biostatistics, Karolinska Institutet and Departments of Clinical Microbiology and Pathology, Karolinska Hospital, Stockholm, Sweden
[c] Clinical Microbiology, Skåne University Hospital Malmö, SE 20502 Malmö, Sweden
[d] Dermatology Clinic, Läkarhuset, Göteborg, Sweden
[e] Department of Dermatology and Venereology, Sahlgrenska University Hospital, Institute of Clinical Sciences at the Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 17 July 2013
Returned to author for revisions
27 August 2013
Accepted 12 September 2013
Available online 10 October 2013

Keywords:
Human papillomaviruses
Skin lesions

ABSTRACT

Pools of frozen biopsies from patients with squamous cell carcinoma (SCC) ($n=29$) actinic keratosis (AK) ($n=31$), keratoacanthoma ($n=91$) and swab samples from 84 SCCs and 91 AKs were analysed with an extended HPV general primer PCR and high-throughput sequencing of amplimers. We found 273 different HPV isolates (87 known HPV types, 139 previously known HPV sequences (putative types) and 47 sequences from novel putative HPV types). Among the new sequences, five clustered in genus Betapapillomavirus and 42 in genus Gammapapillomavirus. Resequencing of the three pools between 21 to 70 times resulted in the detection of 283 different known or putative HPV types, with 156 different sequences found in only one of the pools. Type-specific PCRs for 37 putative types from an additional 296 patients found only two of these putative types. In conclusion, skin lesions contain a large diversity of HPV types, but most appeared to be rare infections.

Introduction

More than 150 different human papillomavirus (HPVs) types have been completely cloned, sequenced and given an official number and the number of putative novel HPV types is continuously growing (Bernard et al., 2010; Bottalico et al., 2011; Chen et al., 2007a; Chen et al., 2007b; Chouhy et al., 2010; Ekstrom et al., 2011; Ekstrom et al., 2010; Foulongne et al., 2012; Kohler et al., 2011; Li et al., 2012; Nobre et al., 2009; Vasiljevic et al., 2008; Vasiljevic et al., 2007). The oncogenic mucosal types of HPV cause cervical cancer (Walboomers et al., 1999), as well as vulvar, anal and penile cancer (IARC, 2007) whereas some cutaneous HPV types cause skin warts and others are associated with development of squamous cell carcinoma (SCC) in patients with the rare immunosuppressive disease epidermodysplasia verruciformis (Jablonska et al., 1972, 1997). Several cutaneous HPV types are commonly found in different skin lesions such as squamous cell carcinoma (SCC) (Asgari et al., 2008; Harwood et al., 2004), actinic keratosis (AK) (Mackintosh et al., 2009), and keratoacanthoma (KA) (Forslund et al., 2003; Stockfleth et al., 1999) as well as on healthy skin (Antonsson et al., 2000; de Koning et al., 2007, 2009; Forslund et al., 2004), in both immunocompetent and immunosuppressive patients.

Classification of papillomaviruses (PVs) is based on the sequence of the major capsid protein gene L1, where the sequence of a new HPV type should be $<90\%$ similar to the L1 gene in any known type (de Villiers et al., 2004). Several general primer PCR systems targeting the L1 gene can amplify a broad range of HPV types (Berkhout et al., 1995; de Roda Husman et al., 1995; Forslund et al., 1999, 2003; Gravitt et al., 2000; Harwood et al., 1999). The general primer pair FAP59/64 amplifies both genital and cutaneous types within the HPV genera Alpha-, Beta- and Gammapapillomavirus (Forslund et al., 1999).

Several different methods for type-specific HPV detection exist (Forslund et al., 1994; Poljak et al., 1999; Schmitt et al., 2006; Soderlund-Strand et al., 2008), notably PCR followed by hybridization to type-specific probes coupled to fluorescent beads (Michael et al., 2011; Schmitt et al., 2011), type-specific PCR and general primer PCR, followed by sequencing (Berkhout et al., 1995; Chouhy et al., 2010; Ekstrom et al., 2010; Forslund et al., 2007).

Previously, we found that high throughput sequencing of amplimers obtained using FAP59/64 revealed an extended diversity of HPV types (Ekstrom et al., 2011). In this study, we used bidirectional sequencing using the recently developed 454 Titanium chemistry to investigate if the detection of HPV sequences

* Correspondence to: Department of Laboratory Medicine, Karolinska Institutet, Huddinge campus F56, Stockholm, Sweden. Fax: +46 40 337 312.
E-mail address: joakim.dillner@ki.se (J. Dillner).

0042-6822/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.virol.2013.09.010

could be improved. In addition, we wished to investigate if the new HPV sequences were commonly present in samples from different skin lesions. The pooled PCR products were purified using MinElute PCR Purification kit (Qiagen), and sequenced using a 454 GS Junior (Roche, Mannheim, Germany).

Results

Identification of known and previously unknown HPV sequences

The analyses of the three different sample pools with different set of primer combinations identified 273 (228 with FAP and 90 with the novel primer combinations) different HPV types or putative types out of which 87 (76 detected with FAP and 33 with novel primer combinations) were known HPV types, 139 (117 detected with FAP and 45 with novel primer combinations) sequences from previously known putative HPV types and 47 (35 with FAP and 12 with novel primer combinations) subgenomic sequences putatively representing novel types (Tables 1 and 2). In a recent report we identified 44 subgenomic sequences from novel putative HPV types, designated SE1 to SE44, of lengths varying from 84 to 450 bp (Ekstrom et al., 2011). In the present study, we used bidirectional sequencing with the GS FLX Titanium chemistry and obtained much longer sequences than in the previous study (mean read length 395 base pairs, compared to 230 base pairs in our previous study). The longer sequences revealed that seven of the previously reported subgenomic SE sequences belonged to the same virus as other SE sequences (SE4=SE27; SE13=SE36; SE17=SE43; SE18=SE29=SE37; SE20=SE44; and SE31=SE33). Sequences SE2, SE6, SE8, and SE21, which were considered as preliminary in the previous publication because of insertion/deletion errors causing premature stop codons in the L1 ORF (Ekstrom et al., 2011), were detected again, but now with more high quality sequences that did not contain premature stop codons. Out of the other 37 SE types detected in the previous study, 33 with FAP and 13 with novel primer combinations were detected again (Table 2). For 17 of these sequences (SE2, 3, 4, 8, 12, 13, 17, 18, 24, 25, 33, 34, 35, 38, 39, 41 and 42), we obtained longer sequences. The longer SE38 sequence was found to overlap with the subgenomic sequence GC04, previously described by Chouhy et al. (2010).

For 33 of the 47 novel putative HPV types detected in this study we obtained >331 bp (range 331–444 bp and mean 436 bp) and for the 14 additional sequences we obtained a sequence from either the 5′-end (n=10, 141–363 bp, and mean=262 bp) or the 3′-end of the amplimer (n=4, 222–225 bp, and mean=224 bp) (Table 1).

The pool with swab samples from SCCs and AKs contained most of the sequences from known/putative HPV types (n=163) closely followed by the biopsies from KAs (known/putative HPV types: n=150), both pools contained similar number of novel putative HPV types (n=24) (Table 1). The majority of the known HPV sequences found belonged to HPV types in genus Gammapapillomavirus (n=128), followed by genus Betapapillomavirus (n=73) and genus Alphapapillomavirus (n=25) (Table 1). Based on the genera of the closest hit in BLAST, most of the novel putative HPV types also belonged to the genus Gammapapillomavirus (n=42), five sequences belonged to genus Betapapillomavirus and none to genus Alphapapillomavirus.

Resequencing

The three sample pools (A, B and C) were resequenced using 121 different multiplex identifiers (MIDs).

An adjusted value for the number of reads from each MID was calculated by multiplying the number of reads for each MID with the number of MIDs used in that particular MID-pool, i.e. 21, 30, 35 and 35. The average adjusted number of reads was 77305 (median 71,820 and range 21,595–280,000). One MID that generated only 3605 reads was considered inadequate and was excluded.

Altogether sequences from 283 different known and putative HPV types were detected. Sample A (fresh frozen biopsies from 29 SCCs and 31 AKs) was resequenced using 30 different MIDs. Sequences from 93 known and putative HPV types were detected, but only three types were detected by all the 30 MIDs (Table 3). Twenty types/putative types were detected by more than 50% of the analyses. Sample B (fresh frozen biopsies from 91 KAs) was resequenced using 70 different MIDs in two separate sequencing runs. In total, 198 known and putative HPV types were detected (Table 3). Sequences from 59 HPV types/putative types were detected by more than half of the resequencing, but only six types were detected in all 70 MIDs. Pool C (swab samples from the top of 84 SCCs and 91 AKs) was resequenced with 21 different MIDs. In

Table 1
Number and genera of HPV sequences found.

Sample	Primers	Known HPV sequences				Novel HPV sequences[a]			Genera of known HPV sequences[d]			Genera of novel putative HPV types[d]	
		HPV types	FA types	FAIMVS	SE types	Complete fragment[a]	Partial, 5′	Partial, 3′	α	β	γ	β	γ
Frozen biopsies from 29 SCCs[b] and 31 AKs[b]	FAP59/FAP64	36	31	3	10	4	1	0	1	47	32	1	4
	Novel primers	13	11	0	0	1	0	0	0	13	11	0	1
Frozen biopsies from 91 KAs[b]	FAP59/FAP64	63	50	7	21	13	6	3	23	43	75	2	20
	Novel primers	12	13	0	3	2	0	0	1	13	14	0	2
Swab samples from 84 SCCs and 91 AKs	FAP59/FAP64	46	62	6	24	9	5	1	2	55	80	1	14
	Novel primers	26	17	2	10	9	0	0	0	35	20	2	7
All samples[c]		87	91	8	40	33	10	4	25	73	128	5	42

[a] >400 bp sequences represent an almost complete or a complete FAP amplimer. Partial sequences may map to either the 5′ or the 3′ of the FAP amplimer.
[b] SCC=squamous cell carcinoma, AK=actinic keratosis, and KA=keratoachantoma.
[c] As sequences were found in more than one of the pools the sum of the sequences is less than the total number of sequences.
[d] Genera based on the top hit sequence in BLAST.

Table 2
A list of HPV sequences found.

Known HPV sequences				Known putative sequences (FA, FAIMVS and SE types)				Novel putative HPV sequences			
Name	Genus	Found in patient group[a]	Found with novel primers	Name	Genus	Found in patient group[a]	Found with novel primers	Name	Genus	Found in patient group[a]	Found with novel primers
HPV 2	Alpha	B		FA1.1	Gamma		C	SE48	Beta	B	
HPV 3	Alpha	B		FA2.1	Gamma	C		SE49	Gamma	A, C	
HPV 4	Gamma	B		FA6	Gamma	A, C	C	SE50	Gamma	B	
HPV 5	Beta	A, B, C	A, B, C	FA8	Gamma	B, C		SE51	Gamma	B, C	
HPV 6	Alpha	B		FA9	Gamma	A, C		SE52	Gamma	B	
HPV 8	Beta	A, C	A, C	FA10	Gamma	B, C		SE53	Gamma	C	
HPV 9	Beta	A, B, C		FA11	Gamma	B	B, C	SE54	Gamma	B	
HPV 10	Alpha	B	B	FA12.2	Gamma	B, C		SE55	Gamma	C	
HPV 11	Alpha	B		FA13	Gamma	A, B, C		SE56	Beta	A	
HPV 12	Beta	A, B, C		FA14	Beta	A, B, C	C	SE57	Gamma	C	
HPV 14	Beta	A, B, C	A, C	FA15	Gamma	B	C	SE58	Gamma	B	
HPV 15	Beta	C	C	FA20.2	Gamma		A	SE59	Gamma	B	
HPV 16	Alpha	B		FA20.3	Gamma	B, C	B	SE62	Gamma	C	
HPV 17	Beta	A, B, C	C	FA23.3	Beta		C	SE63	Gamma	C	
HPV 18	Alpha	B		FA24.2	Gamma	B, C	A	SE64	Gamma	A	
HPV 19	Beta	A, B, C	B	FA25	Beta	B, C		SE65	Gamma	B	
HPV 20	Beta	A, B, C		FA27	Gamma	A, C		SE67	Gamma	C	
HPV 21	Beta	A, B, C		FA28	Gamma	B, C		SE68	Beta	B, C	
HPV 22	Beta	B	A, C	FA31	Gamma	A, B, C		SE69	Gamma	C	
HPV 23	Beta	A, B, C		FA32	Gamma	C		SE70	Gamma	B, C	
HPV 24	Beta		C	FA34	Gamma	A, B, C		SE71	Gamma	B	
HPV 25	Beta		C	FA35	Gamma	A, B, C		SE72	Gamma	B, C	
HPV 27	Alpha	B		FA38	Gamma[b]	C		SE73	Gamma	B	
HPV 28	Alpha	B		FA39	Beta		C	SE74	Gamma	C	
HPV 31	Alpha	B		FA40	Beta	A, B, C		SE75	Gamma	A, B	
HPV 33	Alpha	B		FA41	Gamma	B, C		SE77	Gamma	C	
HPV 35	Alpha	B		FA44	Gamma		B	SE78	Gamma	C	
HPV 36	Beta	A, C		FA45	Gamma[b]	A, B, C		SE79	Gamma	A	
HPV 37	Beta	A, B, C		FA46	Gamma	A, B, C		SE80	Gamma	B	
HPV 38	Beta	A, B, C	B, C	FA48	Gamma	A, C		SE81	Gamma	B	
HPV 45	Alpha	B		FA54	Gamma	B, C		SE82	Gamma	B	
HPV 47	Beta	C		FA57	Gamma	B, C		SE83	Gamma	B	
HPV 48	Gamma		C	FA60.1	Beta	C		SE84	Gamma	B	
HPV 49	Beta	A, B, C		FA61	Gamma	B		SE85	Gamma	B	
HPV 50	Gamma	A, B, C		FA62	Gamma	A	A	SE86	Gamma	B	
HPV 51	Alpha	B		FA64	Gamma	B, C		SE117	Gamma		C
HPV 52	Alpha	B		FA65	Gamma	B, C	A	SE118	Gamma		C
HPV 56	Alpha	B		FA66	Gamma	B	B	SE119	Gamma		C
HPV 57	Alpha	C		FA67	Gamma	A, B, C		SE120	Gamma		C
HPV 58	Alpha	B		FA68	Gamma	B, C		SE121	Gamma		C
HPV 59	Alpha	B		FA69	Gamma	A, B		SE122	Beta		C
HPV 66	Alpha	B		FA70	Beta	A		SE123	Gamma		C
HPV 68	Alpha	C		FA73	Gamma	B		SE124	Gamma		C
HPV 73	Alpha	B		FA77	Gamma[b]	C		SE125	Beta		C
HPV 75	Beta	A		FA81	Gamma	A, B, C		SE126	Gamma		A
HPV 76	Beta	A, B, C		FA86	Gamma	C		SE127	Gamma		B
HPV 77	Alpha	B		FA87	Gamma	C		SE128	Gamma		B
HPV 80	Beta	A, B, C	C	FA88	Gamma	A, C	A				
HPV 82	Alpha	B		FA89	Gamma	A, B, C	B				
HPV 92	Beta	A, B, C		FA90	Gamma	B					
HPV 93	Beta	B, C		FA92	Gamma	B					
HPV 94	Alpha	A, B		FA95	Gamma	C					

HPV 96	Beta	A, C	A
HPV 98	Beta	A, B, C	
HPV 99	Beta	A, C	
HPV 100	Beta	A, B, C	B
HPV 104	Beta	A, B, C	A, C
HPV 105	Beta	A, B, C	A, C
HPV 107	Beta	A, B, C	C
HPV 110	Beta	B, C	C
HPV 111	Beta	A, C	A, C
HPV 115	Beta	B	B, C
HPV 118	Beta	A, B, C	
HPV 119	Gamma	B	
HPV 120	Beta	A, B, C	B, C
HPV 121	Gamma	C	
HPV 122	Beta	A, B, C	B, C
HPV 123	Gamma	B, C	
HPV 124	Beta	A, B, C	A, B, C
HPV 128	Gamma		B, C
HPV 130	Gamma	A, B, C	A, B, C
HPV 131	Gamma	B, C	
HPV 132	Gamma	B, C	
HPV 133	Gamma	B	
HPV 134	Gamma	A, B, C	
HPV 143	Beta		C
HPV 147	Gamma		A, C
HPV 148	Gamma		A, C
HPV 150	Beta	A, B, C	
HPV 151	Beta	A, B, C	
HPV 153	Gamma		C
HPV 156	Gamma		B
RTRX7	Beta	C	
HPV isolate 915 F 06 005 FS1	Gamma	B	
HPV isolate 915 F 06 002 KN1	Gamma		A
HPV isolate 915 F 06 002 KN2	Gamma		A, C
HPV isolate 915 F 06 002 KN3	Gamma	C	

FA100	Gamma	B	
FA101	Gamma	B, C	
FA105	Gamma	B	
FA107.1	Gamma	B	
FA108	Beta	B, C	
FA109	Gamma	A, B, C	
FA111	Gamma	B, C	B
FA112	Beta	A, C	
FA114	Beta	A, C	
FA115	Gamma	B, C	
FA117	Gamma	C	
FA118	Beta	A, B, C	C
FA119	Beta		A, B, C
FA120	Gamma	B	
FA121	Gamma	A, B, C	
FA124	Beta		B, C
FA127	Beta	A, B, C	B
FA126	Beta		A
FA128	Beta	A, C	
FA131	Beta[b]	A, B, C	
FA132	Beta	A, B, C	A, B
FA133	Gamma	C	
FA138	Gamma	B	
FA141	Beta	A, C	
FA142	Gamma	A, C	
FA145	Gamma	C	
FA148	Gamma	C	
FA149	Beta		B, C
FA152	Gamma	C	
FA153	Gamma	C	
FA156	Gamma	C	
FA159	Gamma	B, C	
FA160	Gamma	B, C	B, C
FA162	Beta[b]	A, B, C	C
FA166	Gamma		A
FA168	Beta		A, C
FA170	Gamma		B, C
FA173	Gamma		C
FAIMVS4	Gamma	B, C	
FAIMVS6.2	Beta		C
FAIMVS7	Gamma[b]	A, B, C	
FAIMVS8	Gamma	B	
FAIMVS9	Gamma	B, C	
FAIMVS11.2	Beta[b]	A, B, C	
FAIMVS13.2	Beta[b]	B, C	C
FAIMVS15	Gamma[b]	A, B, C	
SE1	Gamma	C	
SE2	Gamma	C	
SE3	Gamma	A, C	
SE4	Gamma	A, C	
SE5	Gamma	B	
SE6	Gamma	B	
SE7	Gamma	A, B	
SE8	Gamma	B	
SE9	Gamma	B	B
SE10	Gamma	B	

Table 2 (*continued*)

Known HPV sequences				Known putative sequences (FA, FAIMVS and SE types)				Novel putative HPV sequences			
Name	Genus	Found in patient group[a]	Found with novel primers	Name	Genus	Found in patient group[a]	Found with novel primers	Name	Genus	Found in patient group[a]	Found with novel primers
				SE11	Gamma	B, C					
				SE12	Gamma	B, C					
				SE13	Gamma	B, C					
				SE14	Gamma	C					
				SE16	Gamma	A, B, C					
				SE17	Gamma	C					
				SE18	Gamma	A, B, C					
				SE20	Gamma	C					
				SE21	Gamma	A, B, C	B				
				SE22	Beta	B, C	C				
				SE23	Beta	A, B, C					
				SE24	Gamma	C					
				SE25	Gamma	B					
				SE26	Gamma	A, B	C				
				SE28	Gamma	A, C					
				SE30	Gamma	C					
				SE33	Beta	A, C	C				
				SE34	Gamma	B					
				SE35	Gamma	B, C					
				SE38 (GC04)	Gamma	B, C					
				SE39	Gamma	B, C					
				SE41	Gamma	B, C					
				SE42[c]	Gamma	C	C				
				SE56	Beta		C				
				SE57	Gamma		C				
				SE63	Gamma		C				
				SE68	Beta		C				
				SE75	Gamma		C				
				SE80	Gamma		B				
				SE87	Gamma		C				
				Chimpanzee PV CAA8	Beta[b]	B					

[a] A=frozen biopsy from squamous cell carcinomas (SCCs) or actinic keratoses (AKs), B=frozen biopsy from keratoacanthomas and C=swab samples from SCCs or AKs.
[b] Genera based on closest relative with an assigned genus.
[c] SE42 has been cloned and completely sequenced and is now named HPV 155.

Table 3
Repeatable detection of HPV types in the resequencings.

HPV sequences detected in resequencing of Pool A	Proportion of sequencings positive (%)	HPV sequence detected in resequencing of Pool B	Proportion of sequencings positive (%)	HPV sequence detected in resequencing of Pool C	Proportion of sequencings positive (%)
HPV24, HPV108, HPV130	100	HPV38, HPV98, HPV134, FA89, FA111, SE21	100	HPV5, HPV8, HPV9, HPV12, HPV14, HPV19, HPV20, HPV21, HPV23, HPV37, HPV38, HPV49, HPV76, HPV98, HPV104, HPV105, HPV107, HPV120, HPV122, HPV123, HPV124, HPV130, HPV134, HPV136, HPV142, HPV145, HPV146, HPV147, HPV151, HPV isolate cat/EAA/USA/2001, HPV isolate SIBX8, FA14, FA24, FA26, FA34, FA39, FA41, FA65, FA89, FA98, FA118, FA126, FA127, FA141, FA142, FA160, FA162, FAIMVS6, FAIMVS13, FAIMVS15, SE17, SE22, SE23, SE32, SE39, SE40, SE41	100
HPV12, HPV14, HPV23, HPV76, HPV124, FA26, FA37, SE16, SE28	> 75–99	HPV5, HPV9, HPV12, HPV14, HPV19, HPV20, HPV49, HPV80 HPV105, HPV110, HPV120, HPV122, HPV124, HPV133, HPV136, HPV139, HPV147, FA20, FA24, FA41, FA57, FA66, FA81, FA98, FA101, FA107, FA118, FA127, FA132, FA155, FA160, FAIMVS7, FAIMVS15, SE6, SE7, SE13, SE32	> 75–99	HPV17, HPV50, HPV80, HPV96, HPV110, HPV isolate P5-58, FA12, FA32, FA57, FA101, FA132, FAIMVS11, SE4, SE16, SE28, SE38, SE49, SE68	> 75–99
HPV5, HPV9, HPV8, HPV49, HPV50, FAIMVS11, FAIMVS15, SE18	> 50–75	HPV17, HPV21, HPV107, HPV151, FA11, FA34, FA46, FA54, FA65, FA90, FA115, FA126, SE9, SE16, SE25, SE71	> 50–75	HPV24, HPV100, HPV111, HPV132, HPV150, FA3, FA6, FA38, FA121, FA129, FA152, FAIMVS7, SE20, SE33, SE67	> 50–75
HPV25, HPV80, HPV96, HPV120, HPV145, FA1, FA34, FA119, FA127, FA131, FA149, HPV isolate cat/EAA/USA/2001, SE21	> 25–50	HPV4, HPV23, HPV37, HPV50, HPV100, HPV135, HPV138, HPV141, HPV isolate P5-58, FA15, FA18, FA45, FA53, FA72, FA73, FA131, FA159, FAIMVS9, FAIMVS13, SE5, SE10, SE12, SE28, SE34, SE65, SE72, SE86	> 25–50	FA10, FA148, FA153, FA23, FA25, FA45, FA46, FA86, FA9, FAIMVS4, HPV92, SE11, SE24, SE51, SE70	> 25–50
HPV17, HPV19, HPV21, HPV36, HPV38, HPV75, HPV92, HPV94, HPV99, HPV100, HPV104, HPV105, HPV111, HPV122, HPV134, HPV142, HPV147, HPV150, HPV151, HPV isolate B52_FAB, HPV isolate GC05, HPV isolate GC10, HPV isolate GC15, HPV isolate P5-60, FA6, FA9, FA14, FA27, FA35, FA39, FA45, FA46, FA48, FA62, FA70, FA74, FA109, FA112, FA118, FA121, FA124, FA126, FA128, FA132, FA142, FA162, FA81, FA89, FAIMVS7, SE3, SE4, SE23, SE26, SE33, SE40, SE49, SE56, SE64, SE75, SE79	> 0–25	HPV2, HPV3, HPV6, HPV8, HPV10, HPV11, HPV16, HPV18, HPV22, HPV24, HPV27, HPV28, HPV31, HPV33, HPV35, HPV45, HPV51, HPV52, HPV56, HPV58, HPV59, HPV66, HPV68, HPV68, HPV73, HPV76, HPV77, HPV92, HPV96, HPV104, HPV123, HPV130, HPV131, HPV142, HPV146, HPV150, Feline papillomavirus, Chimpanzee papillomavirus isolate CAA8, HPV isolate B2-GP, HPV isolate 915 F 06 005 FS1, HPV isolate BA10 L1, HPV isolate GC12_1, FA1, FA2, FA7, FA9, FA10, FA12, FA14, FA22, FA23, FA25, FA27, FA36, FA40, FA44, FA51, FA61, FA64, FA68, FA69, FA78, FA88, FA92, FA100, FA105, FA108, FA109, FA114, FA116, FA119, FA120, FA121, FA138, FA142, FA158, FA162, FAIMVS4, FAIMVS6, FAIMVS8, SE8, SE11, SE15, SE18, SE22, SE23, SE26, SE35, SE38, SE39, SE41, SE48, SE50, SE51, SE54, SE58, SE59, SE61, SE62, SE66, SE68, SE70, SE73, SE75, SE76, SE80, SE81, SE82, SE83, SE84, SE85	> 0–25	HPV15, HPV24, HPV38, HPV57, HPV68, HPV99, HPV121, HPV131, HPV135, Feline papillomavirus, HPV isolate RTRX7, FA1, FA2, FA8, FA16, FA20, FA27, FA37, FA48, FA51, FA52, FA54, FA60, FA64, FA68, FA77, FA81, FA87, FA88, FA95, FA108, FA109, FA111, FA112, FA117, FA119, FA122, FA128, FA131, FA133, FA134, FA135, FA145, FA150, FA156, FA158, FA159, FAIMVS9, FAIMVS14, FANIMVS17, SE1, SE2, SE3, SE12, SE13, SE21, SE35, SE53, SE55, SE57, SE60, SE61, SE62, SE63, SE69, SE72, SE74, SE77, SE78	> 0–25

total, sequences from 174 HPV-types (including known and putative HPV types) were detected (Table 3) of which 90 were detected by more than half of the resequencings. Sequences from 57 different HPV types/putative types were detected by all the 21 MIDs in the resequencing. In total 55 different sequences were detected in all three pools, 72 different sequences were detected in two of the pools and 156 different sequences in only one of the pools.

Phylogenetic analysis

Phylogenetic relationships between the 73 putative novel HPV types in the "SE-series" (found using 454 sequencing of FAP and novel primer combinations amplimers, 31 of which were reported previously (Ekstrom et al., 2011)), complete, nearly complete fragments or > 200 bp containing the 5′-end, and representative known PV types from all other genera found that 65 of the SE sequences cluster with HPV types of the genus Gammapapillomavirus and eight within the genus Betapapillomavirus (Fig. 1).

Typing of SE sequences using Luminex

Sequence-specific probes for SE sequences 1–44 were used to investigate the prevalence of these sequences in various skin samples. As described above, some of these SE sequences were now found to belong to the same virus. Among the 37 unique SE sequences, 17 were detected in the Luminex analysis of the samples included in the pool where these sequences had originally been detected (Table 4). The SE sequences were mostly detected

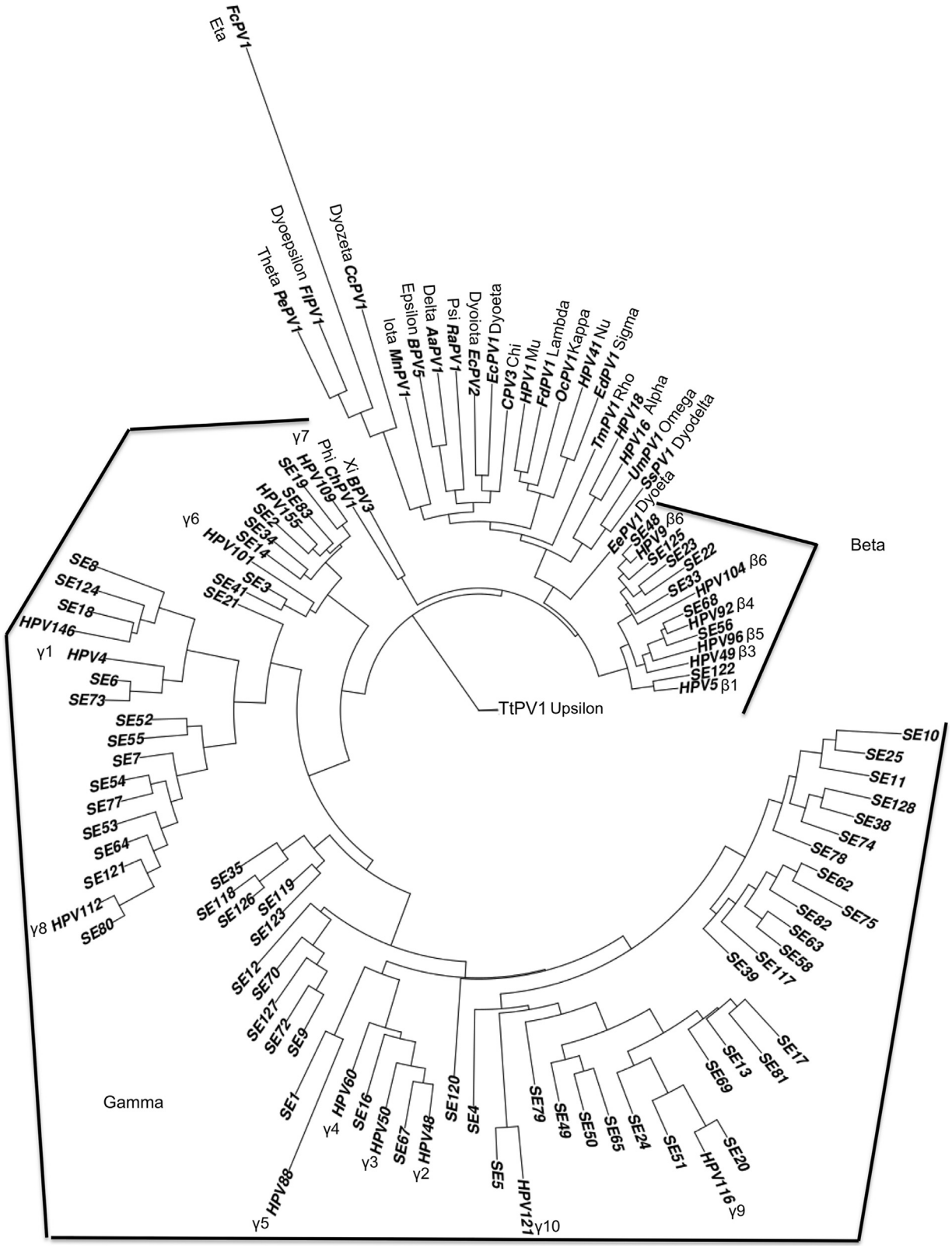

Fig. 1. Maximum likelihood tree based on 42 of the novel putative SE sequences identified in this study with complete or almost complete fragments or >200 bp containing the 5′-end, 31 SE sequences from the previous study (Ekstrom et al., 2011) and L1 from representative known PV types from all other genera.

only in a single sample, the exception being SE41 that was detected in five samples. Fourteen of the 17 SE sequences were found in swab samples, two SE sequences in biopsies, and one SE sequence in both swab samples and a biopsy. For two of the sequences that were corrected in the present paper (SE21 and SE29), there was a mismatch between the corrected template sequence and the probe, which could explain why these sequences were not detected.

Table 4
Results of the Luminex analysis of the presence of the 44 SE sequences in the sample pools where the sequences were originally detected (Ekstrom et al., 2011).

SE type	Luminex-confirmed origin of SE sequence according to diagnosis and sample type
SE9	1 KA biopsy
SE10	1 SCC swab
SE13	1 SCC swab
SE15	1 SCC swab, 1 AK swab
SE17	1 SCC swab, 1 AK swab
SE19	1 SCC swab, 1 AK swab
SE20	1 SCC swab
SE22	1 SCC swab
SE23	1 SCC swab
SE25	1 SCC swab
SE28	1 AK biopsy
SE32	1 SCC swab, 1 AK swab
SE38	1 SCC swab
SE39	1 SCC swab
SE40	1 SCC swab
SE41	2 SCC swab, 2 AK swab, 1 KA biopsy
HPV155 (SE42)	1 SCC swab
SE27, SE29, SE31, SE36, SE37, SE43, SE44	Deleted, found to be identical to other SE sequence.
SE1–SE8, SE11, SE12, SE14, SE16, SE18, SE21, SE24, SE26, SE30, SE33–35	Not detected

Luminex testing of the additional 592 skin samples for the 44 SE sequences found only two sequences (SE16 in one tumour biopsy from a patient with AKs and in one biopsy from normal skin from a BCC patient, and SE41 in one BCC tumour sample).

Discussion

We find that the human skin harbours a large spectrum of different HPV types and that there is reason to believe that a substantial number of HPV types remains to be detected. In this study, we analysed samples that had previously been analysed for the presence of HPV with both conventional and 454 sequencing (Ekstrom et al., 2011; Forslund et al., 2003, 2007, 2004). Nevertheless, we managed to identify 35 subgenomic sequences from novel putative HPV types, which had previously not been detected, by simply using a more modern version of the 454 sequencing method. In order to detect a broader number of HPV types, we also designed novel primer combinations. Two of the novel PCR reactions were specifically designed to detect HPVs form the genera Mu and Nu, which are rather distantly related to other HPV genera and contain only a few types. As all the 12 new putative HPV types detected clustered in the Beta- and Gammapapillomavirus genera, it appears that these genera have a particularly high genomic diversity. Chouhy et al. (2010) have speculated that the HPV general primer system FAP may have reached its limit in detecting new viruses, but our results show that by using updated 454 sequencing, even HPV types present in low copy numbers (that probably would otherwise be missed in the cloning step) can be found. Thus, the high throughput sequencing technology appears to be useful in the continuing efforts to obtain the complete picture of the HPV types that are present in humans.

The mean read-length for 454 GS Junior using titanium chemistry is approximately 400 bases, depending on the sample and genomic characteristics. Using bidirectional sequencing, a complete FA fragment of approximately 440 bp could be obtained. For most of the subgenomic sequences from putative novel types (33/47), the complete or almost complete FA fragment was obtained. The 14 shorter sequences could be the result of DNA breakage or may have been trimmed in the bioinformatics step that removed low quality sequences.

With increasing sensitivity and throughput of the sequencing technology, there is a substantial risk to obtain artifactual "chimeric" sequences that are actually composed of sequences from different viruses. Genomic recombination has been described for cetacean papillomaviruses (Robles-Sikisaka et al., 2012). Even though this has not been described for HPVs, multiple co-infections of related microorganisms can result in recombinations (Liu et al., 2010). Both naturally occurring genomic recombinations and PCR mediated recombinations may mislead phylogenetic analysis (Liu et al., 2010). Our bioinformatics pipeline has been trained to automatically detect and remove putative chimeras, which may have resulted from PCR-mediated recombination between related templates. However, manual checking still found five sequences that were likely chimeras. For all of these five sequences the part of the sequence that appeared to have a different phylogenetic origin was too short to be found by our pipeline. Although we checked the sequences for possible chimeras both automatically and manually, the possibility still exists that some chimeric sequences may have escaped detection. For construction of the phylogenetic tree we used only the sequences validated to not represent chimeras.

The majority of HPV types and putative types detected belonged to the genera Beta- and Gammapapillomavirus, but 25 mucosal types from genus Alphapapillomavirus were also found. Previous studies have also reported on the presence of anogenital oncogenic HPV types in skin samples (Asgari et al., 2008; Iftner et al., 2003). Contamination of the skin by viruses originating from mucosal surfaces, mediated by the fingers has been reported (Alam et al., 2003). We took the biopsies from SCCs and AKs after tape-stripping the skin surface (Forslund et al., 2004), to reduce the probability of detecting contaminating viruses. In the biopsies that had been cleansed with tape-stripping, we found only a single Alphapapillomavirus-type of HPV (HPV 94), suggesting that the presence of HPV94 in this specimen did represent an infection.

Phylogenetically, all the SE sequences were classified into the genera Beta- and Gammapapillomavirus, with the majority (65/73), among the Gammapapillomavirus types. The genus Gammapapillomavirus has been growing rapidly, now comprising at least 58 completely sequenced HPV types and approximately 178 putative HPV types (among them 99 SE sequences) (Bernard et al., 2010; Bottalico et al., 2011; Chen et al., 2007a; Chen et al., 2007b; Ekstrom et al., 2011; Ekstrom et al., 2010; Foulongne et al., 2012; Kohler et al., 2011; Li et al., 2012; Nobre et al., 2009 and unpublished observations). The genus Gammapapillomavirus has a large genetic diversity and many of the SE sequences form clusters outside the previously defined species. Also the genus Betapapillomavirus is a diverse genus with 47 completely sequenced HPV types and approximately 38 putative HPV types (among them 9 SE sequences) (Bernard et al., 2010; Bottalico et al., 2011; Chouhy et al., 2010; Ekstrom et al., 2011; Kohler et al., 2011; Vasiljevic et al., 2008, 2007). Including the sequences reported in this study, there is now evidence to suggest that >300 different cutaneous HPV types exist.

Only 17 of the 44 SE sequences could be detected in the Luminex-based testing of the same samples where these sequences had originally been found by the 454 sequencing. Possible reasons include low copy numbers and/or mismatches to the PCR primer sequences. The proportion of specimens testing positive for these viruses precludes conclusions about their possible disease-association. However, it is noteworthy that although we tested a large number of swabs and biopsies from healthy skin only a single healthy skin biopsy was positive for these viruses. Because design and validation of a large number of type-specific probes for

Luminex is costly and time-consuming and the sensitivity of Luminex detection appeared to be considerably less than the 454 sequencing, we considered that repeat high-throughput sequencing will probably be the best method for continuing epidemiological studies on the role of different HPV types in skin lesions.

The use of multiplex identifiers (MIDs) allows for the separation of individual samples from each other in a single sequencing run. We obtained sequencing data from 121 different MID-adaptors in a resequencing of the three sample pools. The resequencing provided a clear distinction between those HPV types that were detected by all MIDs and are probably present at high copy numbers and those HPV types that were only picked up by a few of the MIDs and are therefore probably present at a much lower abundance. Four of the six SE types that clustered with the Betapapillomavirus types in the phylogenetic tree were detected in more than half of the analyses whereas only 15 of the 55 SE types clustered within the genus Gammapapillomavirus.

In conclusion, we demonstrate that preamplification by general primer PCR followed by bidirectional high throughput sequencing is effective for detection of a very large number of HPV types present in skin samples, including a large number of sequences that had previously have not been described. The continuing improvements of the sequencing technology are likely to continue to reveal an extraordinary and expanding diversity of cutaneous HPVs.

Methods and Materials

Patients

Samples were collected from immunocompetent patients with SCC ($n=119$), AK ($n=114$), basal cell carcinoma (BCC) ($n=117$), keratoacathoma (KA) ($n=8$), seborrhoeic keratosis ($n=46$) and two other benign lesions (one prurigo nodularis and one benign hyperkeratotic skin lesion) attending Swedish and Austrian hospitals (Forslund et al., 2007). Two swab samples from the top of the lesion and from normal adjacent skin were collected by a pre-wetted (0.9% NaCl), cotton-tipped swab that was rolled over the lesion (within margins of the lesion) and suspended in 1 ml of saline. From each lesion as well as from normal adjacent skin, a 2-mm punch-biopsy was taken. The biopsies were taken after tape-stripping (Forslund et al., 2004). Three patients were excluded (one patient was on immunosuppressive drugs, one patient had a mixed tumour and for one patient the diagnosis of SCC could not be confirmed histopathologically). The DNA from the biopsies was extracted using a phenol-free method and the swab samples were extracted by freeze-thawing (Forslund et al., 1999)

An additional 92 biopsies from KAs in both immunosuppressed and immunocompetent patients were collected at the Department of Dermatology and Plastic Surgery at the Norwegian National Hospital in Oslo, Norway. The DNA was extracted using the QIAamp DNA Minikit (Qiagen, Hilden, Germany) (Forslund et al., 2003).

Informed consent was obtained from participants. The study adhered to the declaration of Helsinki and was approved by the Ethical Review Committees of Karolinska Institute and of Lund University (Sweden), Medical University Vienna (Austria) and Institutional Review Board in Oslo (Norway).

Sample preparation

The samples were amplified with the general primer pair FAP, FAP59/64, as described previously (Forslund et al., 1999) using 5 μl of each sample per reaction. For SCC and AK biopsies, only samples that were HPV positive in a previous study (Forslund et al., 2007) were used (SCC, $n=29$ and AK, $n=31$). A plasmid including the sequence of HPV12 was used as a positive control and a sensitivity of one copy per microlitre was found. The PCR amplimers were pooled into three different pools: (A) fresh frozen biopsies from 29 SCC lesions and 31 AK lesions, (B) fresh frozen biopsies from 91 KA lesions and (C) swab samples from the top of the lesion from 84 SCCs and 91 AKs. The PCR products were purified using MinElute PCR Purification kit (Qiagen), pooled and sequenced using a 454 GS Junior (Roche, Mannheim, Germany). Pool A and C were sequenced in two runs each, one without the multiplex identifiers (MIDs) and one using MIDs. Pool B was sequenced in three runs, one without MIDs and two using MIDs. Twelve MIDs (unique oligonucleotide sequence of 10 bases, part of the adaptor sequence) were purchased from Roche and 120 additional MIDs from IDT (Integrated DNA Technology, Leuven, Belgium).

Four DNA libraries were prepared according to the Rapid library preparation method manual for GS Junior Titanium Series (454 sequencing; Roche) using four different pools of MID-adaptors. One pool of 21 MID-adaptors (MID numbers 12–32) was ligated to skin pool C, another pool of 30 MID-adaptors (MID numbers 33–62) was ligated to skin pool A and two pools of 35 MID-adaptors each (MID numbers 63–97 and 98–132 respectively) were ligated to skin pool B. After library preparation, the four libraries with the four different sets of MID-pools were sequenced individually in four 454 GS Junior runs.

In order to investigate whether the broad detection of HPV types by the FAP primers could be further improved, we aligned the FAP primer sequences against all HPV sequences present in GenBank and designed new primers that should be able to detect those HPV types where the FAP primers had mismatches. For the forward primer FAP6085 (Forslund et al., 2003), five additional primers were found to be required to achieve an essentially complete coverage of HPV types. These were FAP6085A CCWGAYCCHAATAARTTTGG (designed to improve the detection of in particular Alphapapillomavirus types), FAP6085B CCWGACCCHAATMRRTTTGC (improved fit to several HPV types in several genera), FAP6085C CCWGATCCHAACMRRTTTGC (improved fit to several HPV types in several genera), FAP6085D GCWGATCCHAATMRRTTTGC (optimal fit to the Mu type 1a), FAP6085E CCWAATCCHAATMCCTTTGC (optimal fit to the Nu type 41). For the reverse primer FAP64 (Forslund et al., 1999), four additional primers were found to be required to achieve adequate coverage of published types FAP64B CCWATGTCWVHCATITCICCATC (improved fit to several HPV types in several genera), FAP64C CCWGTATCWVHCATITCICCATC (improved fit to in particular some Alphapapillomavirus types), FAP64D CCWACATCWVHCATITCICCATC (improved fit to HPV73) and FAP64F CCWGTGTCWVHCATIGTICCGTC (Improved fit to HPV88).

The novel primers were evaluated for ability to amplify different plasmids representing the HPV types where an improved detectability was predicted (HPV1a, 26, 41, 59, 63, 73, 88, 92, 96, 100, 109, 112, and 155). Human DNA from the placenta, 10 ng (3300 copies), was also tested to assure that the primers did not amplify human DNA.

After confirming detection of at least 2 copies/μl of each plasmid, the novel primers were combined together with the original FAP6085 and FAP64 in three different reactions, reaction 1 (multiplex reaction for a broader detection of HPVs, including FAP6085, FAP6085A, FAP6085B, FAP085C, FAP64, FAP64B, FAP64C, FAP64D and FAP64F), reaction 2 (multiplex reaction optimized for detection of HPV types belonging to the genus Mupapillomavirus, including FAP6085C, FAP6805D, FAP64 AND FAP64B) and reaction 3 (reaction optimized for detection of the genus Nupapillomavirus, including FAP6085E and FAP64B).

PCR amplification was then performed on the three pools A, B and C described above using 2 μl DNA in 50 μl reaction containing 1 × Qiagen Multiplex PCR Master Mix Qiagen (Germany) and 0.2 μM of each primer. The thermal programme started with a pre-

heat of 95 °C for 15 min, followed by 45 cycles at 94 °C 30 s, 46–48 °C 90 s, and 72 °C 90 s, with a final extension at 72 °C for 10 min.

PCR products from each reaction were purified and sequenced as described above. Each pool was sequenced separately, using different MIDs for each PCR reaction.

Analysis of sequences

Ambiguous bases (with a Phred quality score less than 30, equivalent to a base calling accuracy of 99.9% (1 error in 1000 bases) (Ewing and Green, 1998)) and primer sequences were removed. Reads with a length of more than 80 bases and less than 20% of ambiguous bases were screened for similarity to human and bacterial DNA using SAHHA2 software (Ning et al., 2001). Reads with at least 95% identity over 75% of their length to human or bacterial DNA were removed. Remaining sequences were processed for assembly of contiguous sequences (contigs) using MIRA software (with the parameters '-job=denovo,genome,accurate,454 454_SETTINGS -AL:mo=40 -AL:mrs=98 -AL:egp=yes -AS:ardct=3 -CO:mrpg=4 -CL:qcmq=30 -CL:bsqcmq=60') (Chevreux et al., 2004). All assembled contigs and singletons were then compared against GenBank using the NCBI BLASTn (reward for nucleotide match=1; penalty of nucleotide mismatch=1; cost to open a gap=0; cost to extend a gap=2; e-value $\leq e^{-4}$). Sequences that over their length had different degrees of similarity to the most closely related sequence in GenBank were considered as "possible chimeras", i.e. artifacts possibly containing sequences originating from different viruses. The sequence that aligned to its most closely related sequence in GenBank was divided into three equal segments. If at least one of the segments had less than 90% similarity and at least one had more than 90% similarity, as well as if the difference between these segments by similarities to corresponding overlapping parts was more than 5% (for example if segment 1 was 88% similar and segment 2 was 94% similar) the sequence was considered as "possible chimera" and excluded from further analysis. If there was less than 5% difference between the above described segments by similarities to corresponding overlapping parts (for example segment 1 was 89% similar and segment 2 was 93% similar) the sequence was checked by manual BLASTn analysis of whether the sequence had an approximately equal frequency of nucleotide matches to its top blast hit across the overlapping part.

If part of the selected sequences ($>=30$ bases) did not overlap to the top hit in the GenBank search, the sequence was cut and blasted again to NCBI nucleotide database. If the top blast hit of these cut sequences was different from the top blast hit of the other part and/or the difference between their per cent identities were $>=10\%$, the sequence was also considered as "possible chimera" and excluded from further analysis. For the resequencing reproducibility evaluation, chimera detection analysis was not performed.

Selected HPV-related sequences that were 90% identical to each other over 90% of their length were clustered using the Cd-hit algorithm (Li and Godzik, 2006) and one representative sequence was selected from each cluster to remove redundancy. HPV-related sequences, which had less than 90% identity over 90% of their length to known HPV genomes, were classified as putatively new HPV types.

Phylogenetic analysis

PhyML v3 (Guindon and Gascuel, 2003) was used to construct a maximum likelihood tree. ModelTest v3.7 (Posada and Crandall, 1998) identified a GTR+I+G substitution model as the best fit model based on the Akaike Information Criterion.

The analyses were restricted to sequences that were complete or almost complete fragments or contained >200 bp of the 5′-of end of the amplimer, to avoid the possibility that non-overlapping sequences might derive from the same virus.

The novel HPV DNA sequences have been deposited in GenBank under the following accession numbers: SE2, JQ250743; SE6, JQ250756; SE8, JQ250773; SE21, JN129832; SE48, JQ250744; SE49, JQ250745; SE50, JQ250746; SE51, JQ250747; SE52, JQ250748; SE53, JQ250749; SE54, JQ250750; SE55, JQ250751; SE56, JQ250752; SE57, JQ250753; SE58, JQ250754; SE59, JQ250755; SE61, JQ250757; SE62, JQ250758; SE63, JQ250759; SE64, JQ250760; SE65, JQ250761; SE67, JQ250762; SE68, JQ250763; SE69, JQ250764; SE70, JQ250765; SE71, JQ250766; SE72, JQ250767; SE73, XXX SE74, JQ250768; SE75, JQ250769; SE77, JQ250770; SE78, JQ250771; SE79, JQ250772; SE80, JX316020; SE81, JX316021; SE82, JX316022; SE83, JX316023; SE84, JX316024; SE85, KC107232; SE86, JX316025; SE117,KC878001; SE118, KC878002; SE119,KC878003; SE120,KC878004; SE121,KC878005; SE122,KC878006; SE123,KC878007; SE124,KC878008; SE125,KC878009; SE126,KC878010; SE127,KC878011 and SE128,KC878012.

Luminex

Unique probes were designed for each of 44 novel subgenomic sequences (SE1-SE44) (Supplementary Table 1). There was no consistent pattern of cross-reactivity with any particular probe when samples were positive for the different viruses, except for the probe for SE44 that had unspecific hybridization and was therefore excluded from analysis. Each probe (DNA technology, Denmark) was covalently linked to fluorescence-labeled carboxy-coated polystyrene beads (BioRad, Sweden) as previously described (Schmitt et al., 2006), with some modifications. In short, 12.5 million carboxylated beads were suspended in 125 μl of 0.1 M 2-(*N*-morpholino) ethanesulfonic acid, pH 4.5 (MES), after which 2 nmol of probes and 1 mg of *N*-(3-dimethylaminopropyl)-*N*-ethyl-carbodiimide (EDC) were added, and the mix was incubated for 30 min under agitation in the dark. After repeated addition of EDC followed by incubation, the beads were first washed with 1.5 ml of 0.02% Tween 20 and then with 1.5 ml of 0.1% sodium dodecyl sulphate before storage in 500 μl of TE buffer at 4 °C. Biotinylated amplicons in 10 μl aliquots from the FAP PCR reaction were hybridized to the probes on the Bioplex 200 Luminex system (Bio-Rad, Sweden), as previously described (Schmitt et al., 2006) but with a hybridization temperature of 50 °C and with a resuspension of beads for 30 min in 70 μl of streptavidin-R-phycoerytrin (Invitrogen, Sweden). Water and 10 ng/μl of human DNA (Sigma, Sweden) in Tris-EDTA buffer were negative controls. Cutoff was determined for each probe by adding the mean of the median fluorescence intensity (MFI) of the water controls to five times the standard deviation of these MFIs, with a minimum cutoff set to three MFIs. All positive samples were re-analysed and only repeatable results were considered positive. In case of discrepancy, the analysis was repeated and the result from two out of three analyses was used.

We first applied the Luminex assay for the 44 putatively new HPV types to 331 samples (fresh frozen biopsies from 30 SCCs, 33 AKs and 92 KAs; swab samples from the top of 84 SCCs and 92 AKs (15)) from the pools where the 44 putatively new types had been detected (12 samples that had been included in the pools were no longer available). In this validation step, the samples in each pool were only tested for the sequences detected in that specific pool. Subsequently, an additional 592 samples from 296 patients were screened in the Luminex assay (fresh frozen biopsies from 67 SCCs, 117 BCCs, 56 AKs, 8 KAs and 48 benign lesions as well as a healthy skin biopsy from each patient) for all sequences detected in the three pools.

Samples diluted 1:2 in TE buffer (10 mM Tris–HCl, 0.1 mM EDTA, pH 8.0) were amplified using FAP primers (Cybergene, Sweden) as described, using a 5 μl input volume from each sample (Forslund et al., 1999). All forward primers were biotinylated at the 5′ end.

Acknowledgments

We are grateful for the support from BioCARE, the Swedish Cancer Society and the Swedish Research Council. We thank Kia Sjölin, Camilla Lagheden and Birgitta Möller for excellent technical assistance and Reinhard Kirnbauer, Bernt Lindelöf, Eva Hradil and Ole Petter Clausen for patient enrolment.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2013.09.010.

References

Alam, M., Caldwell, J.B., Eliezri, Y.D., 2003. Human papillomavirus-associated digital squamous cell carcinoma: literature review and report of 21 new cases. Journal of the American Academy of Dermatology 48 (3), 385–393.

Antonsson, A., Forslund, O., Ekberg, H., Sterner, G., Hansson, B.G., 2000. The ubiquity and impressive genomic diversity of human skin papillomaviruses suggest a commensalic nature of these viruses. Journal of Virology 74 (24), 11636–11641.

Asgari, M.M., Kiviat, N.B., Critchlow, C.W., Stern, J.E., Argenyi, Z.B., Raugi, G.J., Berg, D., Odland, P.B., Hawes, S.E., de Villiers, E.M., 2008. Detection of human papillomavirus DNA in cutaneous squamous cell carcinoma among immunocompetent individuals. Journal of Investigative Dermatology 128 (6), 1409–1417.

Berkhout, R.J., Tieben, L.M., Smits, H.L., Bavinck, J.N., Vermeer, B.J., ter Schegget, J., 1995. Nested PCR approach for detection and typing of epidermodysplasia verruciformis-associated human papillomavirus types in cutaneous cancers from renal transplant recipients. Journal of Clinical Microbiology 33 (3), 690–695.

Bernard, H.U., Burk, R.D., Chen, Z., van Doorslaer, K., Hausen, H., de Villiers, E.M., 2010. Classification of papillomaviruses (PVs) based on 189PV types and proposal of taxonomic amendments. Virology 401 (1), 70–79.

Bottalico, D., Chen, Z., Dunne, A., Ostoloza, J., McKinney, S., Sun, C., Schlecht, N.F., Fatahzadeh, M., Herrero, R., Schiffman, M., Burk, R.D., 2011. The oral cavity contains abundant known and novel human papillomaviruses from the Betapapillomavirus and Gammapapillomavirus genera. Journal of Infectious Diseases 204 (5), 787–792.

Chen, Z., Schiffman, M., Herrero, R., Burk, R.D., 2007a. Identification and characterization of two novel human papillomaviruses (HPVs) by overlapping PCR: HPV102 and HPV106. Journal of General Virology 88 (Pt 11), 2952–2955.

Chen, Z., Schiffman, M., Herrero, R., Desalle, R., Burk, R.D., 2007b. Human papillomavirus (HPV) types 101 and 103 isolated from cervicovaginal cells lack an E6 open reading frame (ORF) and are related to gamma-papillomaviruses. Virology 360 (2), 447–453.

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Muller, W.E., Wetter, T., Suhai, S., 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Research 14 (6), 1147–1159.

Chouhy, D., Gorosito, M., Sanchez, A., Serra, E.C., Bergero, A., Fernandez Bussy, R., Giri, A.A., 2010. New generic primer system targeting mucosal/genital and cutaneous human papillomaviruses leads to the characterization of HPV 115, a novel Beta-papillomavirus species 3. Virology 397 (1), 205–216.

de Koning, M.N., Struijk, L., Bavinck, J.N., Kleter, B., ter Schegget, J., Quint, W.G., Feltkamp, M.C., 2007. eta papillomaviruses frequently persist in the skin of healthy individuals. Journal of General Virology 88 (Pt 5), 1489–1495.

de Koning, M.N., Weissenborn, S.J., Abeni, D., Bouwes Bavinck, J.N., Euvrard, S., Green, A.C., Harwood, C.A., Naldi, L., Neale, R., Nindl, I., Proby, C.M., Quint, W.G., Sampogna, F., ter Schegget, J., Struijk, L., Wieland, U., Pfister, H.J., Feltkamp, M.C., 2009. Prevalence and associated factors of betapapillomavirus infections in individuals without cutaneous squamous cell carcinoma. Journal of General Virology 90 (Pt 7), 1611–1621.

de Roda Husman, A.M., Walboomers, J.M., van den Brule, A.J., Meijer, C.J., Snijders, P.J., 1995. The use of general primers GP5 and GP6 elongated at their 3′ ends with adjacent highly conserved sequences improves human papillomavirus detection by PCR. Journal of General Virology 76 (Pt 4), 1057–1062.

de Villiers, E.M., Fauquet, C., Broker, T.R., Bernard, H.U., zur Hausen, H., 2004. Classification of papillomaviruses. Virology 324 (1), 17–27.

Ekstrom, J., Bzhalava, D., Svenback, D., Forslund, O., Dillner, J., 2011. High throughput sequencing reveals diversity of human papillomaviruses in cutaneous lesions. International Journal of Cancer 129 (11), 2643–2650.

Ekstrom, J., Forslund, O., Dillner, J., 2010. Three novel papillomaviruses (HPV109, HPV112 and HPV114) and their presence in cutaneous and mucosal samples. Virology 397 (2), 331–336.

Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Research 8 (3), 186–194.

Forslund, O., Antonsson, A., Nordin, P., Stenquist, B., Hansson, B.G., 1999. A broad range of human papillomavirus types detected with a general PCR method suitable for analysis of cutaneous tumours and normal skin. Journal of General Virology 80 (Pt 9), 2437–2443.

Forslund, O., DeAngelis, P.M., Beigi, M., Schjolberg, A.R., Clausen, O.P., 2003. Identification of human papillomavirus in keratoacanthomas. Journal of Cutaneous Pathology 30 (7), 423–429.

Forslund, O., Hansson, B.G., Bjerre, B., 1994. Typing of human papillomaviruses by consensus polymerase chain reaction and a non-radioactive reverse dot blot hybridization. Journal of Virological Methods 49 (2), 129–139.

Forslund, O., Iftner, T., Andersson, K., Lindelof, B., Hradil, E., Nordin, P., Stenquist, B., Kirnbauer, R., Dillner, J., de Villiers, E.M., 2007. Cutaneous human papillomaviruses found in sun-exposed skin: beta-papillomavirus species 2 predominates in squamous cell carcinoma. Journal of Infectious Diseases 196 (6), 876–883.

Forslund, O., Lindelof, B., Hradil, E., Nordin, P., Stenquist, B., Kirnbauer, R., Slupetzky, K., Dillner, J., 2004. High prevalence of cutaneous human papillomavirus DNA on the top of skin tumors but not in "Stripped" biopsies from the same tumors. Journal of Investigative Dermatology 123 (2), 388–394.

Forslund, O., Ly, H., Higgins, G., 2003. Improved detection of cutaneous human papillomavirus DNA by single tube nested 'hanging droplet' PCR. Journal of Virological Methods 110 (2), 129–136.

Foulongne, V., Sauvage, V., Hebert, C., Dereure, O., Cheval, J., Gouilh, M.A., Pariente, K., Segondy, M., Burguiere, A., Manuguerra, J.C., Caro, V., Eloit, M., 2012. Human skin microbiota: high diversity of DNA viruses identified on the human skin by high throughput sequencing. PLoS ONE 7 (6), e38499.

Gravitt, P.E., Peyton, C.L., Alessi, T.Q., Wheeler, C.M., Coutlee, F., Hildesheim, A., Schiffman, M.H., Scott, D.R., Apple, R.J., 2000. Improved amplification of genital human papillomaviruses. Journal of Clinical Microbiology 38 (1), 357–361.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology 52 (5), 696–704.

Harwood, C.A., Spink, P.J., Surentheran, T., Leigh, I.M., de Villiers, E.M., McGregor, J.M., Proby, C.M., Breuer, J., 1999. Degenerate and nested PCR: a highly sensitive and specific method for detection of human papillomavirus infection in cutaneous warts. Journal of Clinical Microbiology 37 (11), 3545–3555.

Harwood, C.A., Surentheran, T., Sasieni, P., Proby, C.M., Bordea, C., Leigh, I.M., Wojnarowska, F., Breuer, J., McGregor, J.M., 2004. Increased risk of skin cancer associated with the presence of epidermodysplasia verruciformis human papillomavirus types in normal skin. British Journal of Dermatology 150 (5), 949–957.

IARC, 2007. Human papillomaviruses. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans 90, 1–636.

Iftner, A., Klug, S.J., Garbe, C., Blum, A., Stancu, A., Wilczynski, S.P., Iftner, T., 2003. The prevalence of human papillomavirus genotypes in nonmelanoma skin cancers of nonimmunosuppressed individuals identifies high-risk genital types as possible risk factors. Cancer Research 63 (21), 7515–7519.

Jablonska, S., Dabrowski, J., Jakubowicz, K., 1972. Epidermodysplasia verruciformis as a model in studies on the role of papovaviruses in oncogenesis. Cancer Research 32 (3), 583–589.

Jablonska, S., Majewski, S., Obalek, S., Orth, G., 1997. Cutaneous warts. Clinical Dermatology 15 (3), 309–319.

Kohler, A., Gottschling, M., Manning, K., Lehmann, M.D., Schulz, E., Kruger-Corcoran, D., Stockfleth, E., Nindl, I., 2011. Genomic characterization of ten novel cutaneous human papillomaviruses from keratotic lesions of immunosuppressed patients. Journal of General Virology 92 (Pt 7), 1585–1594.

Li, J., Cai, H., Xu, Z., Wang, Q., Hang, D., Shen, N., Liu, M., Zhang, C., Abliz, A., Ke, Y., 2012. Nine complete genome sequences of cutaneous human papillomavirus genotypes isolated from healthy skin of individuals living in rural Henan province, China. Journal of Virology 86 (21), 11936.

Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22 (13), 1658–1659.

Liu, W., Li, Y., Learn, G.H., Rudicell, R.S., Robertson, J.D., Keele, B.F., Ndjango, J.B., Sanz, C. M., Morgan, D.B., Locatelli, S., Gonder, M.K., Kranzusch, P.J., Walsh, P.D., Delaporte, E., Mpoudi-Ngole, E., Georgiev, A.V., Muller, M.N., Shaw, G.M., Peeters, M., Sharp, P. M., Rayner, J.C., Hahn, B.H., 2010. Origin of the human malaria parasite *Plasmodium falciparum* in gorillas. Nature 467 (7314), 420–425.

Mackintosh, L.J., de Koning, M.N., Quint, W.G., Ter Schegget, J., Morgan, I.M., Herd, R.M., Campo, M.S., 2009. Presence of beta human papillomaviruses in nonmelanoma skin cancer from organ transplant recipients and immunocompetent patients in the West of Scotland. British Journal of Dermatology 161 (1), 56–62.

Michael, K.M., Forslund, O., Bacevskij, O., Waterboer, T., Bravo, I.G., Pawlita, M., Schmitt, M., 2011. Bead-based multiplex genotyping of 58 cutaneous human papillomavirus types. Journal of Clinical Microbiology 49 (10), 3560–3567.

Ning, Z., Cox, A.J., Mullikin, J.C., 2001. SSAHA: a fast search method for large DNA databases. Genome Research 11 (10), 1725–1729.

Nobre, R.J., Herraez-Hernandez, E., Fei, J.W., Langbein, L., Kaden, S., Grone, H.J., de Villiers, E.M., 2009. E7 oncoprotein of novel human papillomavirus type 108 lacking the E6 gene induces dysplasia in organotypic keratinocyte cultures. Journal of Virology 83 (7), 2907–2916.

Poljak, M., Brencic, A., Seme, K., Vince, A., Marin, I.J., 1999. Comparative evaluation of first- and second-generation digene hybrid capture assays for detection of human papillomaviruses associated with high or intermediate risk for cervical cancer. Journal of Clinical Microbiology 37 (3), 796–797.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Robles-Sikisaka, R., Rivera, R., Nollens St., H.H., Leger, J., Durden, W.N., Stolen, M., Burchell, J., Wellehan Jr., J.F., 2012. Evidence of recombination and positive selection in cetacean papillomaviruses. Virology 427 (2), 189–197.

Schmitt, M., Bravo, I.G., Snijders, P.J., Gissmann, L., Pawlita, M., Waterboer, T., 2006. Bead-based multiplex genotyping of human papillomaviruses. Journal of Clinical Microbiology 44 (2), 504–512.

Schmitt, M., de Koning, M.N., Eekhof, J.A., Quint, W.G., Pawlita, M., 2011. Evaluation of a novel multiplex human papillomavirus (HPV) genotyping assay for HPV types in skin warts. Journal of Clinical Microbiology 49 (9), 3262–3267.

Soderlund-Strand, A., Dillner, J., Carlson, J., 2008. High-throughput genotyping of oncogenic human papilloma viruses with MALDI-TOF mass spectrometry. Clinical Chemistry 54 (1), 86–92.

Stockfleth, E., Meinke, B., Arndt, R., Christophers, E., Meyer, T., 1999. Identification of DNA sequences of both genital and cutaneous HPV types in a small number of keratoacanthomas of nonimmunosuppressed patients. Dermatology 198 (2), 122–125.

Vasiljevic, N., Hazard, K., Dillner, J., Forslund, O., 2008. Four novel human betapapillomaviruses of species 2 preferentially found in actinic keratosis. Journal of General Virology 89 (Pt 10), 2467–2474.

Vasiljevic, N., Hazard, K., Eliasson, L., Ly, H., Hunziker, A., de Villiers, E.M., Norrild, B., Dillner, J., Forslund, O., 2007. Characterization of two novel cutaneous human papillomaviruses, HPV93 and HPV96. Journal of General Virology 88 (Pt 5), 1479–1483.

Walboomers, J.M., Jacobs, M.V., Manos, M.M., Bosch, F.X., Kummer, J.A., Shah, K.V., Snijders, P.J., Peto, J., Meijer, C.J., Munoz, N., 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. Journal of Pathology 189 (1), 12–19.